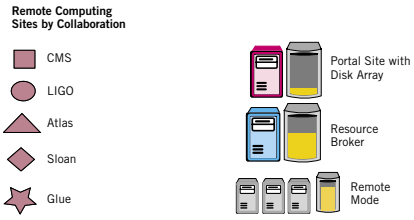
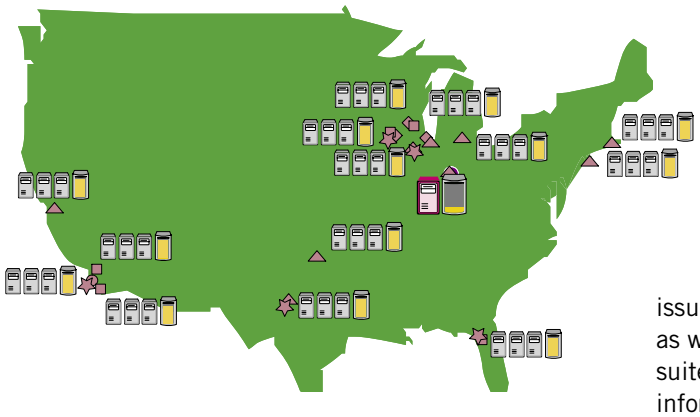
6

International Collaborations

Global Grids are the goal, and we are on our way...





WorldGrid—an intense short-term collaboration between Trillium and the European Grid projects, EDG and DataTAG—is a concrete step toward building world-wide Grid infrastructure.

WorldGrid hopes to learn how to support physicists working collaboratively from their different institutions in the U.S. and in Europe. To prove the Grid really works, we are using production releases of application software from the CMS, ATLAS, LIGO and SDSS collaborations. The goals are to create an integrated Grid spanning the Atlantic, involving many computing sites running either VDT or EDG grid middleware. Underneath a simple web interface, scientists will gain access to the combined EU-US WorldGrid testbed of over 20 sites.

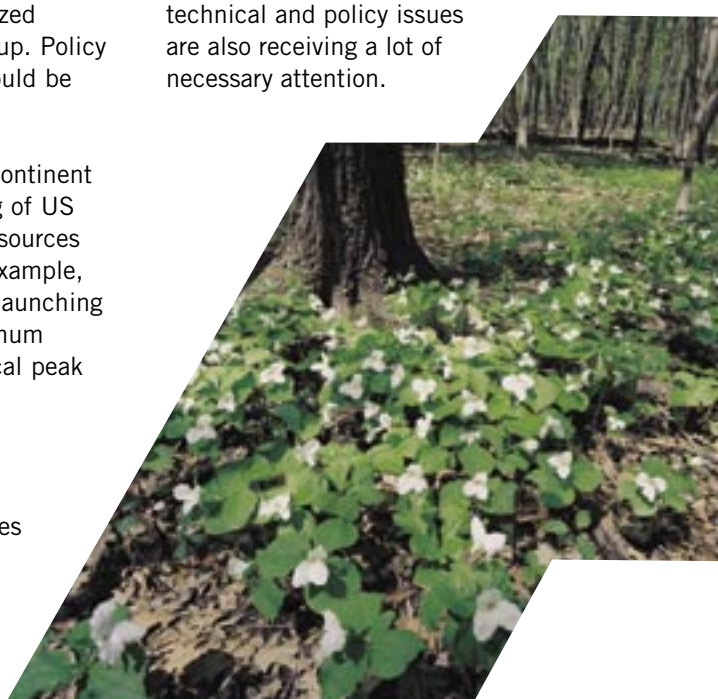
With this trial WorldGrid infrastructure, we are exploring mechanisms for controlled use of shared resources among the experiments. Any experiment may "own" or have rights to use physical and intellectual resources that cross many traditional boundaries: departmental, institutional, regional and national. In order to achieve their scientific mission, members or groups need access to resources spread across these boundaries, subject to agreed policy and security mechanisms.

The policy mechanisms could be viewed through different "renderings" of the WorldGrid usage. Two monitoring systems, Ganglia- and Nagios-based, were developed to collect information on Grid resource usage both in machine and job domains. Information was collected in a hierarchical, scalable fashion, and organized according to experiment group. Policy goals and resulting usage could be compared.

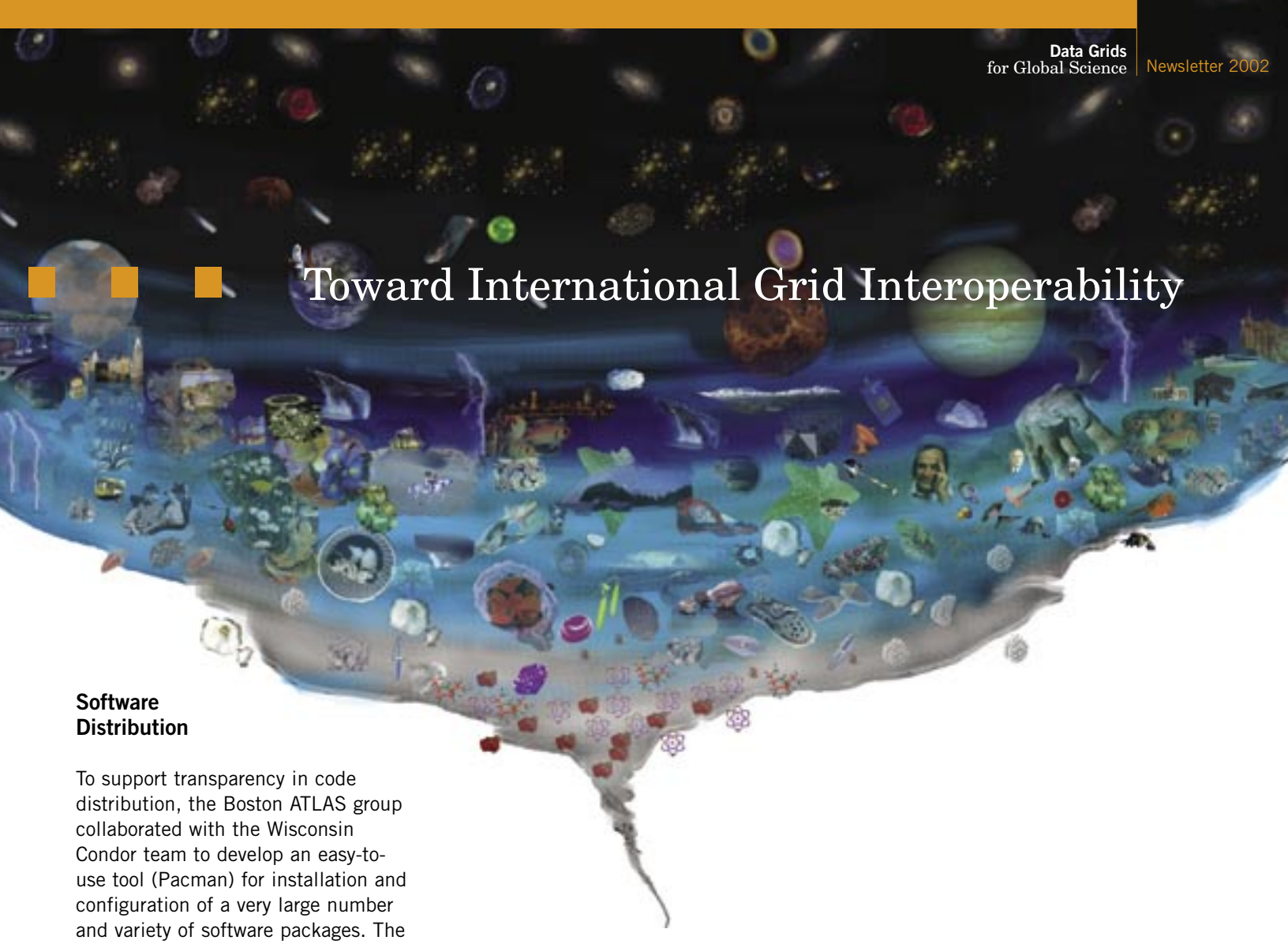
Thus, scientists from each continent can benefit from the sharing of US and EU-based computing resources in different time zones-for example, exploiting the possibility of launching their simulations with maximum performance even during local peak hours.

Grid Laboratory Uniform Environment (GLUE)

The GLUE sub-project focuses on a "bottoms up" approach to interoperability and has a goal to address detailed issues in the grid middleware layers, as well as develop test and regression suites. Consistent resource discovery information is one of the keys to allowing jobs to be scheduled across international grid sites. The GLUE schema have been developed and included as an option in the Globus MDS suite. Cross-authentication and authorization technical and policy issues are also receiving a lot of necessary attention.



Toward International Grid Interoperability



Software Distribution

To support transparency in code distribution, the Boston ATLAS group collaborated with the Wisconsin Condor team to develop an easy-to-use tool (Pacman) for installation and configuration of a very large number and variety of software packages. The Pacman tool is being used to create "installations" which enable new sites to download the VDT, application-specific software components, setup scripts, and specialized grid tool kits with simple commands such as:

```
%pacman -get iVDGL:WorldGrid
        %pacman -get BU:Atlas
        %pacman -get UF:CMS
```

The commands would install not only the grid middleware but also the applications and configuration information needed to run the physics analysis and simulation codes. Pacman is being adopted by other experiments and is being evaluated by the peer European Data Grid projects.

ATLAS Distributed Simulation

The US ATLAS Grid Testbed collaboration has developed a number of Grid tools to facilitate job submission and monitoring on a number of sites. Two programs were developed for job submission. The first, a flexible script-based toolkit, GRATS, was developed at UT Arlington and Oklahoma for rapid testing and production running. A second user interface, GRAPPA, was developed at Indiana, Chicago and Argonne using web service portal technology provided by the computer science department at Indiana University. The portal was first developed to allow job submission to the US ATLAS Testbed, but then expanded to include sites from other experiments (LIGO, SDSS, CMS) and the EU Grids, EDG and DataTAG. Common job submission, resource monitoring, and data management functionality was provided by the portal system to manage physics simulations on both US and EU Grid Testbed, speaking the middleware language of each. Data collections from the common submission were cataloged using the MAGDA replica and metadata management system developed by Brookhaven Laboratory.

The exercise demonstrates essential interoperability between different grid software deployments under the jurisdiction of different software suites. On the US-ATLAS Grid, the jobs were executed on Linux RedHat 7.2 platforms using VDT middleware. In the European DataGrid and DataTAG testbeds, jobs were submitted to a user interface machine in the US which automatically dispatched them to computing resources in Europe via resource brokers at CERN and Pisa. Data collections and monitoring data were returned to the portal system. The job execution could proceed in two modes, either using site-resident software or a more portable method, in which the executable was sent to the computing resource as part of the job. The latter mode provided significant agility, since no assumption was made as to the availability or release level of the code at the target compute facility. Such processes are essential if we are to realize the full power of data grids.